



Heriot-Watt University
Research Gateway

Gaze and Attention During an HRI Storytelling Task

Citation for published version:

Sabyruly, Y, Broz, F, Keller, I & Lohan, KS 2015, 'Gaze and Attention During an HRI Storytelling Task', Artificial Intelligence for Human-Robot Interaction, Arlington, VA, United States, 12/11/15 - 14/11/15.

Link:

[Link to publication record in Heriot-Watt Research Portal](#)

Document Version:

Peer reviewed version

General rights

Copyright for the publications made accessible via Heriot-Watt Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

Heriot-Watt University has made every reasonable effort to ensure that the content in Heriot-Watt Research Portal complies with UK legislation. If you believe that the public display of this file breaches copyright please contact open.access@hw.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Gaze and Attention During an HRI Storytelling Task

Yerlikzhan Sabyruly¹ and Frank Broz¹ and Ingo Keller¹ and Katrin S. Lohan^{1,2 * †}

Introduction

In this paper we discuss the results of an HRI study concerning how an attention system can influence the users perception of a small humanoid (NAO) robot. The participants in this study engage in a storytelling task. We use semi-automated analysis to investigate the change in cognitive load of our participants while they are talking to the robot.

User Study

In this user study, participants were asked to tell a story to a robot using visual aids. The robot used a semiautomatic attention system to respond to the participants in a lifelike manner while they were interacting with it.

Participants were sat opposite a NAO robot (Robotics 2013) and asked to perform a task that involved both storytelling and object assembly. While they performed this task, we collected data using a variety of sensors on the participant, the robot, and within the experimental environment.

Our participants were asked to tell a story about the escape of a thief called Bob to our robot NAO. They were given the story to read in advance and asked to retell the story to the NAO in their own words. They were asked to assemble 3 lego figures, Bob and the 2 police men who were chasing him. They also were supplied with the car Bob was using to escape and the police car which was following him. Finally, they were asked to use a provided lego bridge to illustrate how Bob managed to escape from the police.

Setup

The experimental setup can be seen in Figure 1. To record our sessions we used:

- The NAO camera to record the interaction from the robots perspective, the states of the robot's behaviour were recorded in to log files,
- TobiiGlasses 2 to record the participant's gazing behaviour, pupil diameter and the view from the participants perspective

^{*1}Department of Mathematical and Computer Science, at the Heriot-Watt University, EH14 4AS Edinburgh, United Kingdom

^{†2}k.lohan@hw.ac.uk

Copyright © 2015, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

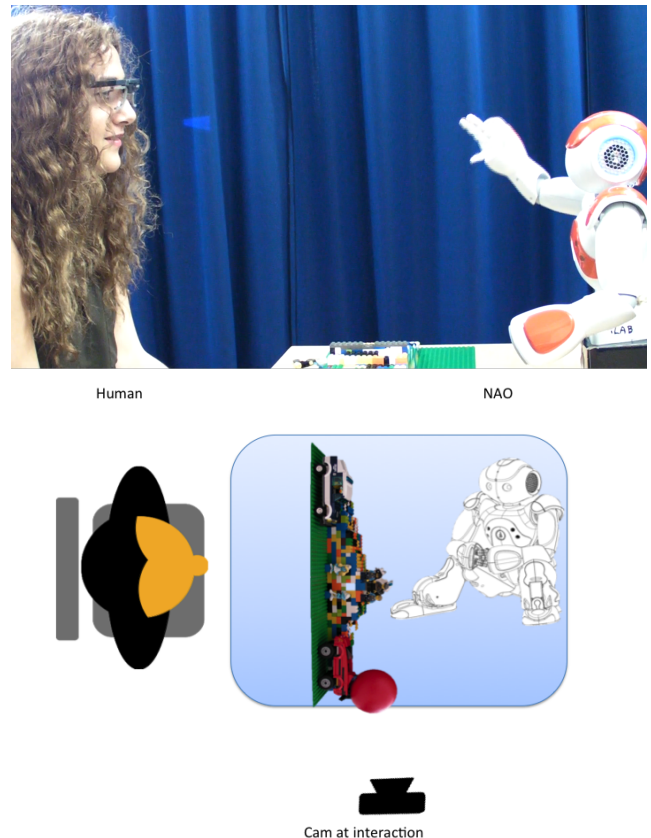


Figure 1: The experimental setup.

- An external camera overlooking the interaction.

Participants were asked to say "hello" to the robot after they were sitting down across of the table from the NAO. They were then asked to use the visual aids to tell the story to the NAO. After they interacted with the NAO, the participants were asked to fill out a survey including a big 5 personality questionnaire.

Participants

All participants were recruited from the university and where either students or staff.

Age group	18-24 years	24-34 years	34-54 years
Participants	12	6	1
Gender	7 F, 4 M, 1 O	5 M, 2 F	1 M

Table 1: The age and gender distribution of our participants.

The Attention System

The semiautomatic attention system which was used in this user study exhibited some remote controlled behaviours and some automated behaviour. The NAO was able to follow the red car and the face of the user. The importance of both objects as well as the interaction partner has been suggested in other attention systems e.g. (Lohan et al. 2012). Furthermore, the NAO could speak to the participant by using a text to speech module. The experimenter would talk to the participant by using the text to speech mechanism during the session if necessary. At the beginning of the session, the robot would follow the gaze of the participant until the red car would appear in the story, then the robot would follow the red car until it is out of its camera field. The robot would greet and say goodbye to the participant in a pre-scripted behaviour, waving and talking.

Measurements

For our measurements, we used the recordings of the Tobii-Glasses 2. They supply video with audio from the user perspective, as well as files comprising the user's gaze in pixel coordinates and pupil diameter. We use this data to calculate the percentage of time the participants spent looking at the NAO, the minimum pupil diameter and the maximum pupil diameter in percent of time.

Data Analysis

To analyse our data we used Python scripts for data processing. For the calculation of when the participants looked at the NAO, we used the SIFT + FLANN key-point detection implementation of the openCV library for python (Budiharto 2014) in order to locate the NAO in the camera image from the TobiiGlasses (Bulling and Gellersen 2010). This allowed the user's gaze data to be processed to determine whether it fell within or without a bounding box around the NAO when the NAO was detected in the image (Figure ??). This method of automated analysis of robot-directed attention using gaze-tracking data is similar to the data analysis performed in other work on visual attention in HRI such as in Yu, Schermerhorn, and Scheutz, for example (Yu, Schermerhorn, and Scheutz 2012). Manual annotation using ELAN was used to mark the beginning and end of the experiment session (from when our participants were sitting down in front of the NAO to when they got up in the end).

To calculate the amount of time when pupil diameter is minimummaximum during looking at NAO we used both the pupil diameter provided by the TobiiGlasses2 .tsv files and the NAO detection.

We also measured the participants experiences with the NAO using a questionnaire.

Results

Questionnaire Results

At the end of the user study each participant filled out the questionnaire. The questionnaire consists of 35 questions. 26 questions are used to determine the participant's big 5 personality traits. The additional 9 questions measure the participant's impression of the robot.

The graph below demonstrates age comparison with how well a person gets on with the robot. The graph shows that the younger people seem to cope with the robot better, however it is only a very slight difference.

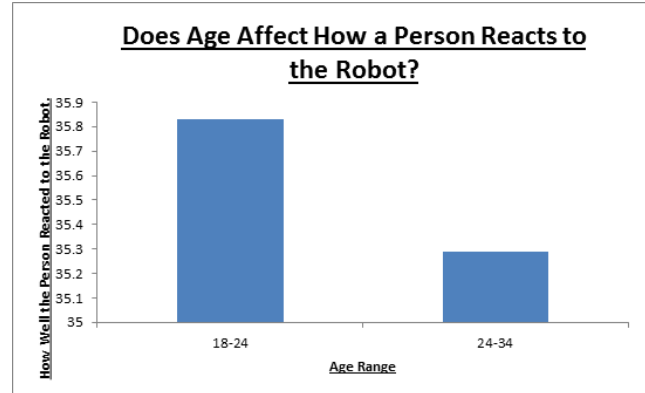


Figure 2: Age comparison with how well person get on with the robot.

The results demonstrated that people who are good with technology are generally better with the robot, but again it was only slightly difference in comparison with people who are not very good with technology. Generally both groups get on well with the robot.

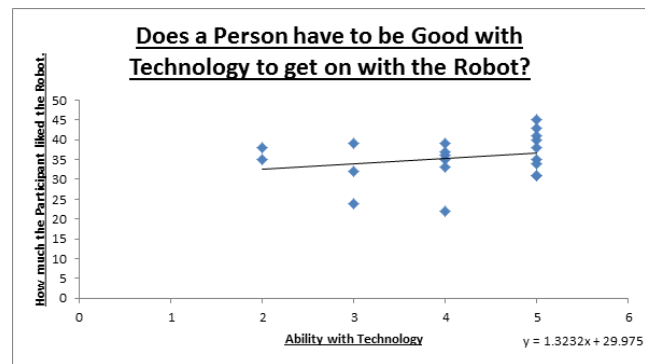


Figure 3: How well people who good with technology with the robot.

The following graph demonstrates whether gender has any impact on how a person reacts to the robot. It appear

that females in our user study get on better with the robot than males do.

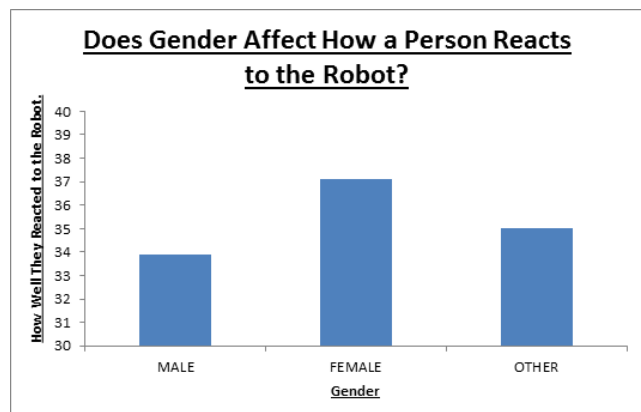


Figure 4: What gender deal better with the robot.

It was also interesting to consider how much time people who liked the robot spent with it. This graph shows some unexpected results. When people liked the robot more, they spent less time with it. One possible explanation for this behavior is that those who were more comfortable with the robot were less nervous and therefore managed to complete the tasks faster.

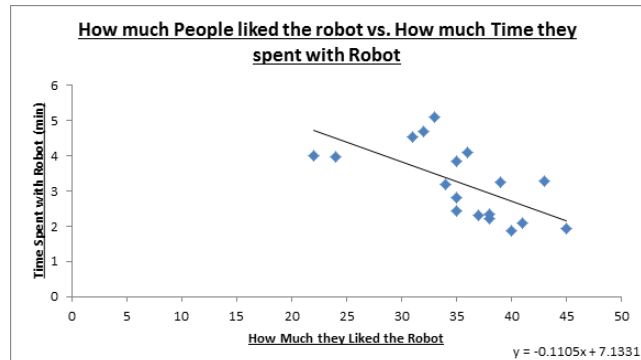


Figure 5: Time spent with the robot.

The next graph illustrates how extraverted a person is against how well they got on with the robot. There does seem to be a correlation showing that extraverts are more comfortable with the robot than introverts are. Perhaps this is because they are more outgoing and friendly. However, the more introverted participants did say that having three cameras on them made them feel uncomfortable.

Finally it is interesting to see the overall impression from the robot. This table shows that only 2 participants have not been satisfied enough, and rated the robot lower than 50 percent on a scale from 1 to 50. The overall rate is 35.4 which

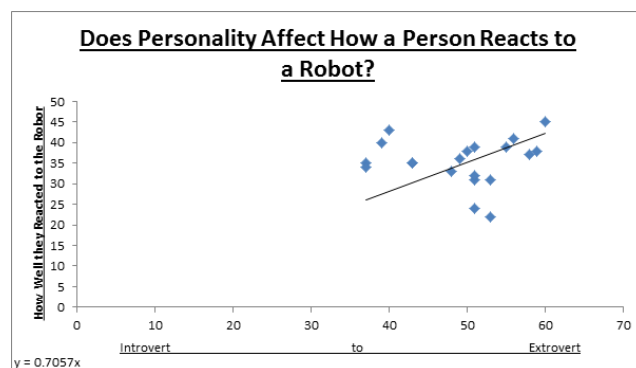


Figure 6: Time spent with the robot.

is good. These ratings suggest that the visual attention system of the robot wasn't bad, but perhaps not good enough. It is obvious that many improvements could be made, and this user study helped to reveal weaknesses in the current system.

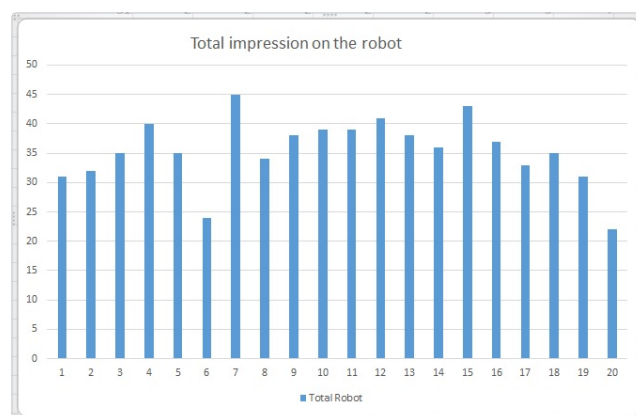


Figure 7: The rates of all participants.

Gaze Data Results

For the camera images from the users' glasses in which the NAO could be automatically detected, the user was looking at the NAO roughly 50% of the time (see figure 8). However, for certain users the mean fell well below this level, suggesting that there may be individual differences contributing to how often users looked at the NAO while telling the story. Notice that this mean is not the same as the mean of how often users looked at the NAO overall during storytelling, as there were times when users looked down entirely at their assembly task and therefore the NAO could not be found in the camera image from the glasses.

One of the goals of this experiment was to investigate the cognitive load that users experienced while performing this storytelling task and how it effected their behaviour towards the robot.

In figure 9, the maximal pupil diameter while the participants are looking at the NAO can be seen. The minimal pupil

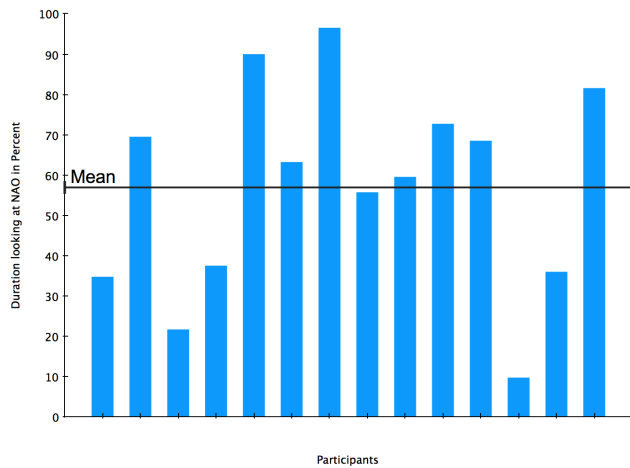


Figure 8: Percentage of time looking at Nao

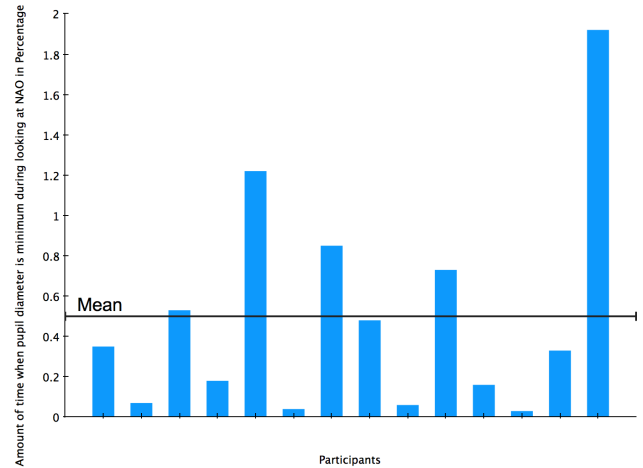


Figure 10: Minimum pupil diameter

diameter while the participants are looking at the NAO can be seen in figure 10.

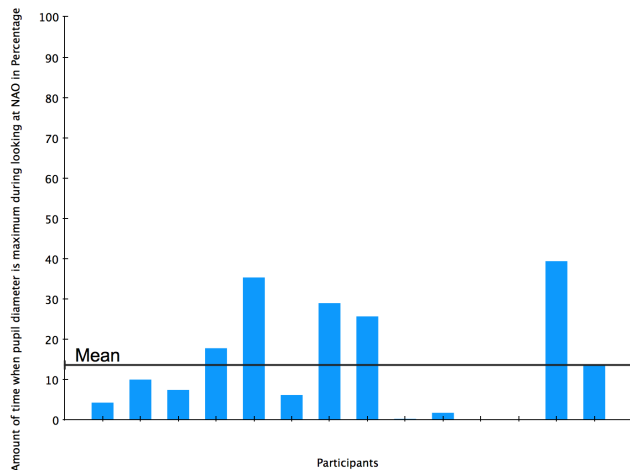


Figure 9: Maximum pupil diameter

The, maximal pupil diameter correlates according to Palinko et al. (Palinko and Sciutti) with the cognitive load of a person. As can be seen in our results the percentage of time the pupil are dialated maximal while they are looking at the NAO is quiet low (under 15%), but the minimal dialation in comparison is almost non existent (below 2 %).

One could assume, that while the participants are telling the story their cognitive load is the highest. We are in the process of analysing the participants' speech to confirm this hypothesis.

Acknowledgments

The authors would like to thank Sharon Macleod for her help, as experimenter. We also appreciate the financial support provided by Heriot-Watt University.

References

- Budiharto, W. 2014. Robust vision-based detection and grasping object for manipulator using sift keypoint detector. In *Advanced Mechatronic Systems (ICAMEchS), 2014 International Conference on*, 448–452. IEEE.
- Bulling, A., and Gellersen, H. 2010. Toward mobile eye-based human-computer interaction. *Pervasive Computing, IEEE* 9(4):8–12.
- Lohan, K. S.; Rohlfing, K.; Pitsch, K.; Saunders, J.; Lehmann, H.; Nehaniv, C.; Fischer, K.; and Wrede, B. 2012. Tutor spotter: Proposing a feature set and evaluating it in a robotic system. *International Journal of Social Robotics* 4:131–146.
- Palinko, O., and Sciutti, A. Exploring the estimation of cognitive load in human robot interaction.
- Robotics, A. 2013. The nao robot. Available at: <http://lrvnw.aldebaran-robotics.com/en/>. (Accessed September 4 2013).
- Yu, C.; Schermerhorn, P.; and Scheutz, M. 2012. Adaptive eye gaze patterns in interactions with human and artificial agents. *ACM Transactions on Interactive Intelligent Systems* 1(2):13.